

In the format provided by the authors and unedited.

# SRSF7 maintains its homeostasis through the expression of Split-ORFs and nuclear body assembly

Vanessa Königs<sup>1,7</sup> , Camila de Oliveira Freitas Machado<sup>1,7</sup>, Benjamin Arnold<sup>1,7</sup> , Nicole Blümel<sup>1,7</sup>, Anfisa Solovyeva<sup>1</sup>, Sinah Löbbert<sup>1</sup>, Michal Schafranek<sup>1</sup>, Igor Ruiz De Los Mozos<sup>2,3</sup> , Ilka Wittig<sup>4</sup>, Francois McNicoll<sup>1</sup> , Marcel H. Schulz<sup>5,6</sup>  and Michaela Müller-McNicoll<sup>1</sup> 

---

<sup>1</sup>Institute of Cell Biology and Neuroscience, Goethe University, Frankfurt am Main, Germany. <sup>2</sup>The Francis Crick Institute, London, UK. <sup>3</sup>Department of Neuromuscular Disease, UCL Institute of Neurology, London, UK. <sup>4</sup>Functional Proteomics Group, Centre of Biochemistry, Medical School, Goethe University, Frankfurt am Main, Germany. <sup>5</sup>Institute of Cardiovascular Regeneration, Medical School, Goethe University, Frankfurt am Main, Germany. <sup>6</sup>German Centre for Cardiovascular Research, Partner site RheinMain, Frankfurt am Main, Germany. <sup>7</sup>These authors contributed equally: Vanessa Königs, Camila de Oliveira Freitas Machado, Benjamin Arnold, Nicole Blümel. ✉e-mail: [mcnicoll@bio.uni-frankfurt.de](mailto:mcnicoll@bio.uni-frankfurt.de); [mueller-mcnicoll@bio.uni-frankfurt.de](mailto:mueller-mcnicoll@bio.uni-frankfurt.de)

## Supplementary material

### Supplementary tables

**Table S1: SRSF7 binds strongly to its own pre-mRNA.** Counts of normalized significant crosslink events (X-links) that map to the entire mouse *SRSF7* gene or separately to the indicated segments quantified from SRSF3- and SRSF7-GFP iCLIP datasets without OE.

X-links	SRSF7	SRSF3
all	1202024	1644147
ORF	9899	218
5'UTR	1234	24
3'UTR	858	111
Introns	19041	400
<i>SRSF7</i> gene	31032	753
Percentage (%)	2.58	0.05

**Table S2: The *SRSF7*-PCE isoform is translated.** Spearman correlation coefficients (Rs) of Ribo-Seq replicates, the number of Ribo-Seq reads that map specifically to the *SRSF7*-PCE isoform, and junction reads mapping to the junction between exon 3 and the PCE.

Sample	Correlation Repl (Rs)	Reads <i>SRSF7</i> -PCE	Junction Reads PCE
WT	0.987	6707	0
<i>SRSF7</i> OE	0.992	22246	8

**Table S3: Comparison piCLIP vs. iCLIP.** Numbers of crosslink events (X-links) quantified from iCLIP and piCLIP data (pooled monosomal and polysomal fractions of 2 replicates each) that map to the entire mouse *SRSF7* gene or separately to the PCE.

X-links	iCLIP	piCLIP
all	4077052	105753
<i>SRSF7</i>	35549	1708
PCE	11913	643
Percentage (%)	33.5	37.6

**Table S4: Binding of SRSF7 variants on the *SRSF7* gene.** Numbers of total crosslink events (X-links) and X-links that map to the *SRSF7* gene in pooled iCLIP samples (2 replicates).

Protein	Significant X-links (FDR<0.05)	X-links on <i>SRSF7</i> gene (%)
SRSF7 endo (WT)	24590	1.09
SRSF7 endo (OE)	24803	12.95
SRSF7_RRM	5707	15.54
SRSF7-GFP	57239	14.43
noAb	31	
noUV	27	

**Table S5: Exon 5-6 junction reads.** Quantification of junction reads from RNA-Seq data to estimate the expression levels of the four *SRSF7* isoforms in P19 cells.

Isoform	WT	(%)	OE	(%)
Full-length	124	47.15	959	50.63
$\Delta$ YFQ	118	44.87	740	39.07
$\Delta$ 11aa	19	7.22	180	9.50
$\Delta$ 27aa	2	0.76	15	0.79
<b>Total</b>	<b>263</b>		<b>1894</b>	

**Table S6: Numbers of X-links and significant binding sites of pooled iCLIP replicates and controls.**

Protein	Unique X-links	Sign. X-links (FDR<0.05)	Sign. binding sites
SRSF7 (OE)	166214511	4232166	247028
mutZn	17800599	4244001	305302
$\Delta$ 27aa	735538	85119	11485
noUV	34027	1251	404

**Table S7. Human NMD targets of RNA-binding proteins with encoded Split-ORFs**

**Table S8: Most enriched protein domains within Split-ORFs (human).**

<b>PFAM_ID</b>	<b>Domain</b>	<b>Number</b>	<b>odds ratio</b>	<b>P-value</b>	<b>FDR</b>
PF00400	WD40 repeat	49	2.7563	0.0000	0.0000
PF13768	Von-Willebrand-factor A	7	9.3710	0.0001	0.0204
PF00076	RNA recognition motif	35	2.0286	0.0004	0.0543
PF16212	P-type ATPase	7	6.6934	0.0004	0.0543
PF16300	WD40 repeat	6	8.0322	0.0005	0.0610
PF00035	dsRNA binding motif	7	5.5123	0.0010	0.0911
PF01302	CAP-Gly domain	6	6.6934	0.0010	0.0911
PF00412	Zinc finger, LIM-type	15	2.7138	0.0012	0.0919
PF00105	Zinc finger, C4-type	12	3.1502	0.0012	0.0919
PF00191	Annexin-like	6	6.1786	0.0014	0.0919

**Table S9. Protein domains within Split-ORFs (human and mouse)**

**Table S10. Ribo-seq read counts in unique NMD exons of RBPs (human and mouse)**

**Table S12. Primers used in this study**

Primer	Sequence (5' → 3')	T <sub>m</sub> [°C]	Purpose
<i>SRSF7</i> exon1 fw	CATCATGTCACGCTACGGGC	69	RT-PCR
<i>SRSF7</i> exon8 rev	ACGGGTGAACTTGAGAACTTCAG	66	RT-PCR
<i>SRSF7</i> exon2 fw	ATGCAGTTCGAGGATTGGAT	63	RT-PCR
<i>SRSF7</i> intron5 fw	AGGGAGGAGAGACCCAATTCCCC	76	RT-PCR
<i>SRSF7</i> exon6 rev	ATCTTGATCTCGACCTTGAG	58	RT-PCR
<i>SRSF7</i> 3'UTR rev	TGGTTGATGATGATCCTTCCTAGGTT	65	RT-PCR
<i>GFP</i> rev	ACTTCCAGGCTCGACGAACC	69	RT-PCR
<i>SRSF3</i> exon1 fw	TGCATCGTGATTCCCTGTCCC	68	RT-PCR
<i>SRSF3</i> exon7 rev	CCTTTCATTTGACCTAGATCGG	64	RT-PCR
<i>CalR</i> fw	TCAAGTCCGGGACAATCTTTG	63	RT-PCR
<i>CalR</i> rev	TCCTCTTTACGCTTCTTGTCCCTC	63	RT-PCR
<i>BA</i> _spliced_fw	GAGCACAGCTTCTTTGCAGCTC	63	RT-PCR
<i>BA</i> _spliced_rev	CTGGGTCATCTTTTCACGGTTG	63	RT-PCR
<i>BA</i> _unspliced_fw	AAGCCTGGGGTTTTCTTGGG	63	RT-PCR
<i>GAPDH</i> fw	GTTTGTGATGGGTGTGAACCACGAG	63	qPCR
<i>GAPDH</i> rev	GAGTGGGAGTTGCTGTTGAGGTCGC	63	qPCR
<i>U6</i> _fw	GCTCGCTTCGGCAGC	63	qPCR
<i>U6</i> _rev	AAATATGGAACGCTTCACGAATT	63	qPCR
<i>SRSF7</i> Intron3 fw	AAGAGGGTCGACCTGCAAAC	63	qPCR
<i>SRSF7</i> Intron3 rev	AGTGTTTGTGTACCTCTGCTGAA	63	qPCR
<i>SRSF7</i> PCE short fw	CAAGCCTCATCTGGCTCTTTTG	63	qPCR
<i>SRSF7</i> PCE short rev	CTAGATCGTGACCTCAAGGTCTA	63	qPCR
<i>SRSF7</i> PCE long fw	CAAGCCTCATCTGGCTCTTTTG	63	qPCR
<i>SRSF7</i> PCE long rev	CTAGATCGTGACCTTGCAAGTTT	63	qPCR
3'RACE RT rev	CCAGTGAGCAGAGTGACGAGGACTGAGCT CAAGCTTTTTTTTTTTTTTTTTTTTTTTT	82	3'RACE
3'RACE PCR rev	CCAGTGAGCAGAGTGACGAGGACTCGAGC TCAAGC	63	3'RACE
<i>SRSF7</i> OE <i>NheI</i> fw	GCGCTAGCATGTCACGCTACGGGCGGT	72	cloning
<i>SRSF7</i> OE <i>KpnI</i> rev	GCGGTACCGCGTCCATTCTTTCTGGACT	70	cloning
<i>SRSF7</i> _RRM <i>KpnI</i> rev	GCGGTACCAGGCGACTGACTCAAACGAAG	70	cloning